

基于视觉感知与学习的图像质量评价

丰明坤¹, 周红², 孙丽慧¹

(1. 浙江科技学院 信息与电子工程学院, 杭州 310023; 2. 浙江省文成县公安局, 浙江 温州 325300)

摘要: 针对几种经典评价方法的缺点,通过引入人眼的主观特性,提出一种基于视觉感知与学习(visual perception and learning, VPL)的方法,以解决人类视觉系统(human vision system, HVS)多通道评价融合的不稳定性。借助反向传播(back propagation, BP)神经网络构建了评价融合模型,分别对几种视觉感知算法的多通道评价进行融合,并基于回归函数对视觉感知算法的结果进行二次互补融合。结果表明,相对于现有主流方法,本文所提方法的各项评价指标均具有较大的优势。

关键词: 图像质量评价;视觉感知;多通道评价;自适应融合

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 1671-8798(2019)06-0444-06

Image quality assessment based on visual perception and learning

FENG Mingkun¹, ZHOU Hong², SUN Lihui¹

(1. School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China; 2. Wencheng County Public Security Bureau, Wenzhou 325300, Zhejiang, China)

Abstract: Aiming at the shortcoming of several classical assessment methods, a new method of image quality assessment based on visual perception and learning(VPL), was proposed by introducing subjective characteristics of human eyes with a view to solving instability of multi-channel assessment pooling of the human vision system(HVS). Then, an assessment pooling model was constructed by virtue of the back propagation (BP) neural network, pooling multi-channel assessment of several visual perception algorithms. Finally, second complementary pooling was conducted for results of every visual perception algorithm based on the regression function. The experiment results show that every assessment indicator of the proposed method has greater advantages compared to the existing prevailing methods.

Keywords: image quality assessment; visual perception; multi-channel assessment; adaptive pooling

收稿日期: 2018-11-05

基金项目: 浙江省公益技术应用研究项目(LGF18F020010)

通信作者: 丰明坤(1978—),男,河南省延津人,讲师,博士,主要从事信息智能感知处理研究。E-mail: 104027@zust.edu.cn。

随着大数据和云计算等业务的兴起,图像信号获得了广泛的应用^[1],因而基于计算机视觉原理的图像质量客观评价引起了人们的重视^[2]。峰值信噪比(peak signal noise ratio,PSNR)是基于像素误差统计结果的经典算法,但其指标水平较低。近些年人们结合工程学和图像特征提出了一些新的评价算法,例如,结构相似度(structural similarity,SSIM)^[3]是基于局部像素域相似度统计结果的工程方法,梯度相似度(gradient similarity mechanism,GSM)^[4]则针对图像梯度特征进行相似度评价,而奇异值分解(singular value decomposition,SVD)^[5]通过提取图像的奇异特征值进行误差评价。上述4种算法从评价准则和特征提取两方面构成了图像质量客观评价方法的基础,于是一些研究者针对这4种算法提出了改进的评价方法^[6-9],但仍存在一些问题:各评价指标水平的差别较大,很难同时兼顾综合性能;应用场景不同,各评价指标水平会出现较大的波动,稳定性较差。针对上述问题,本研究通过引入人类视觉系统(human vision system,HVS)显著感知特性提出了基于视觉感知与学习(visual perception and learning,VPL)的图像质量融合评价方法,即先构建反向传播(back propagation,BP)神经网络融合模型对各算法的多通道评价结果进行融合,再基于回归函数对各算法的评价结果进行二次互补融合。

1 视觉感知与学习评价 VPL 方法的原理

本研究提出的 VPL 方法的评价原理可用图 1 来描述,它充分利用了 HVS 视觉感知特性和视觉处理特性,首先对图像特征进行视觉处理;然后对 PSNR、SSIM、GSM 和 SVD 4 种算法进行改进,提出了对应的视觉峰值信噪比(VPSNR)、视觉结构相似度(VSSIM)、视觉梯度相似度(VGSM)和视觉奇异值分解(VSVD)4 种评价算法;最后基于 BP 神经网络构建多通道评价融合模型,进而获得最终 V_{VPL} 评价分值。VPL 方法包括如下步骤:对参考图像 $r(i,j)$ 和失真图像 $d(i,j)$ 进行高斯卷积处理^[3],模拟 HVS 瞳孔感光原理, (i,j) 代表图像坐标点;基于二维 Log-Gabor 小波变换算法^[10]提取 $r(i,j)$ 和 $d(i,j)$ 的多通道视图 $r_{s,o}(i,j)$ 和 $d_{s,o}(i,j)$,模拟 HVS 亮度感知非线性和信息多通道感知特性,参数 s,o 分别为

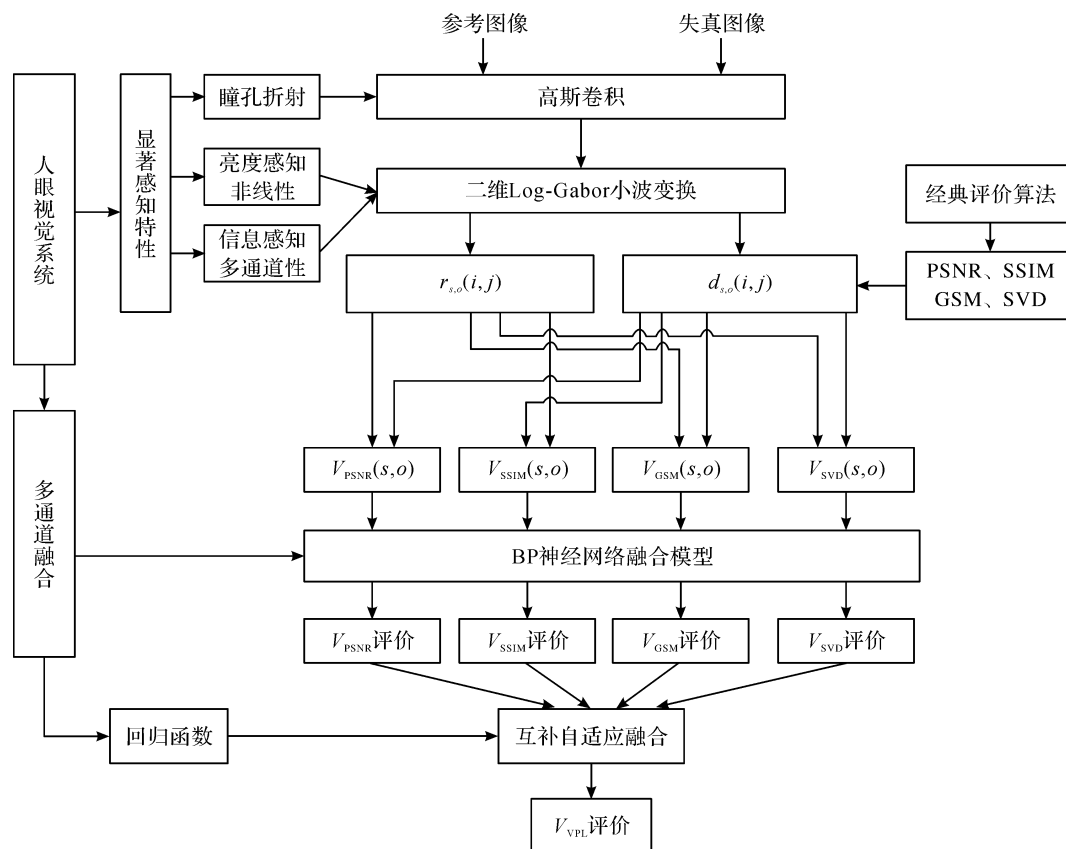


图 1 VPL 方法原理

Fig. 1 Theory of VPL method

Log-Gabor 的尺度因子和方向因子;设计 VPSNR、VSSIM、VGSM 和 VSVD 4 种改进的视觉评价算法,分别对 $d_{s,o}(i,j)$ 进行测量,获得视觉评价 $V_{\text{PSNR}}(s,o)$ 、 $V_{\text{SSIM}}(s,o)$ 、 $V_{\text{GSM}}(s,o)$ 和 $V_{\text{SVD}}(s,o)$;构建 BP 神经网络融合模型,分别通过 $V_{\text{PSNR}}(s,o)$ 、 $V_{\text{SSIM}}(s,o)$ 、 $V_{\text{GSM}}(s,o)$ 和 $V_{\text{SVD}}(s,o)$ 训练集的学习,依次获得对应测试集的融合结果 V_{PSNR} 评价、 V_{SSIM} 评价、 V_{GSM} 评价和 V_{SVD} 评价;基于回归函数对所得结果进行逐步自适应融合,最终获得 $d(i,j)$ 的 V_{VPL} 评价分值。

1.1 视觉感知评价算法原理

1.1.1 视觉峰值信噪比算法(VPSNR)

基于 PSNR 算法及其改进算法^[10],VPSNR 评价公式如下:

$$V_{\text{PSNR}}(s,o) = 10\log_{10}\left(\frac{255^2}{V_{\text{mse}}(s,o)}\right)。(1)$$

式(1)中: $V_{\text{mse}}(s,o)$ 为 $r_{s,o}(i,j)$ 和 $d_{s,o}(i,j)$ 的视觉均方误差,其定义如下:

$$V_{\text{mse}}(s,o) = \frac{1}{L} \sum_{p=1}^L \frac{1}{MN} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \omega(i,j) [r_{s,o}(i,j) - d_{s,o}(i,j)]^2。(2)$$

式(2)中: $\omega(i,j)$ 为局部分块的高斯卷积核函数; MN 为 $\omega(i,j)$ 的窗口尺寸, L 为 $\omega(i,j)$ 的子块数量。

1.1.2 视觉结构相似度算法(VSSIM)

基于 SSIM 算法^[3]及其改进算法^[11],VSSIM 评价公式如下:

$$V_{\text{SSIM}}(s,o) = \frac{1}{L} \sum_{p=1}^L [H_p(s,o)]^{\alpha_1} \times [\zeta_p(s,o)]^{\alpha_2} \times [\psi_p(s,o)]^{\alpha_3}。(3)$$

式(3)中:一般取权重参数 $\alpha_1 = \alpha_2 = \alpha_3 = 1$, $H_p(s,o)$ 、 $\zeta_p(s,o)$ 、 $\psi_p(s,o)$ 的定义分别为

$$\begin{cases} H_p(s,o) = \frac{2\mu_{r,p}(s,o)\mu_{d,p}(s,o) + C_1}{\mu_{r,p}^2(s,o) + \mu_{d,p}^2(s,o) + C_1}; \\ \zeta_p(s,o) = \frac{2\sigma_{r,p}(s,o)\sigma_{d,p}(s,o) + C_2}{\sigma_{r,p}^2(s,o) + \sigma_{d,p}^2(s,o) + C_2}; \\ \psi_p(s,o) = \frac{2\sigma_{r,d,p}(s,o) + C_3}{\sigma_{r,p}(s,o)\sigma_{d,p}(s,o) + C_3}。 \end{cases} (4)$$

式(4)中: C_1 、 C_2 、 C_3 为稳定算法的常数; $\mu_{r,p}(s,o)$ 、 $\mu_{d,p}(s,o)$ 为局部视觉高斯滤波均值; $\sigma_{r,p}(s,o)$ 、 $\sigma_{d,p}(s,o)$ 为局部视觉高斯滤波均方差; $\sigma_{r,d,p}(s,o)$ 为局部视觉高斯滤波协方差,相关定义参见文献[7]。

1.1.3 视觉感知梯度相似度算法(VGSM)

基于 GSM 算法^[4]及其改进算法^[12],VGSM 评价公式如下:

$$\begin{cases} V_{\text{GSM}}(s,o) = \frac{1}{L} \sum_{p=1}^L V_{\text{GSM},p}(s,o); \\ V_{\text{GSM},p}(s,o) = \frac{1}{MN} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \omega(i,j) \frac{2g_{r(s,o)}(i,j)g_{d(s,o)}(i,j) + T}{g_{r(s,o)}^2(i,j) + g_{d(s,o)}^2(i,j) + T}。 \end{cases} (5)$$

式(5)中: $V_{\text{GSM},p}(s,o)$ 为局部视觉高斯滤波的梯度相似度评价; $g_{r(s,o)}(i,j)$ 、 $g_{d(s,o)}(i,j)$ 分别为 $r_{s,o}(i,j)$ 和 $d_{s,o}(i,j)$ 的梯度特征视图,梯度算子参见文献[4]; T 为稳定算法的常数。

1.1.4 视觉感知奇异值分解算法(VSVD)

基于 SVD 算法^[5]及其改进算法^[13],所提 VSVD 评价公式如下:

$$\begin{cases} V_{\text{SVD}}(s,o) = \frac{1}{L} \sum_{p=1}^L [V_{\text{SVD},p}(s,o) - D_{\text{mid}}(s,o)]; \\ V_{\text{SVD},p}(s,o) = \left\{ \frac{1}{Z} \sum_{t=1}^Z [v_{r,t}(s,o) - v_{d,t}(s,o)]^2 \right\}^{\frac{1}{2}}。 \end{cases} (6)$$

式(6)中: $V_{\text{SVD},p}(s,o)$ 为局部视觉高斯滤波的奇异值特征评价; $D_{\text{mid}}(s,o)$ 为 $V_{\text{SVD},p}(s,o)$ 的中间值; $v_{r,t}(s,o)$ 、 $v_{d,t}(s,o)$ 分别为 $r_{s,o}(i,j)$ 和 $d_{s,o}(i,j)$ 的第 t 个奇异特征值; Z 为 $\omega(i,j)$ 窗口内的奇异特征值数量。

1.2 多通道评价融合原理

HVS 多通道融合对不同评价算法表现出不同的参数特性,目前尚未有相关的研究成果揭示 HVS 的

多通道融合机制。本研究借助 BP 神经网络解决 HVS 多通道融合的不稳定性问题,构建了相应的 HVS 多通道融合模型,依据图像失真类型的不同将数据库中的失真图像数量随机对半分为训练集和测试集两部分。BP 神经网络融合模型为图 2 所示的典型三层结构^[14],其中,输入层神经元数量为通道数 $s \times o = 5 \times 4 = 20$,隐藏层神经元数量为 20,输出层神经元数量为 1,以学习误差 $e \leq 0.01$ 和迭代次数 $\rho = 200$ 作为该模型的训练终止判决条件。BP 融合模型的训练阶段首先选取一种图像失真类型和一种视觉评价算法(以 VPSNR 为例),将训练集中所有失真图像的 $V_{\text{PSNR}}(s, o)$ 分值作为模型的训练输入,以差异均值主观分(difference mean opinion score, DMOS)为训练目标,对模型进行训练,当达到其中一个训练终止条件 $e \leq 0.01$ 或 $\rho = 200$ 时终止训练,转入测试阶段。BP 融合模型的测试阶段选取对应训练阶段同一失真类型图像的 $V_{\text{PSNR}}(s, o)$ 分值作为模型的输入,模型的输出即为 $V_{\text{PSNR}}(s, o)$ 的融合结果分值 V_{PSNR} 。依次选取不同的图像失真类型和不同的评价算法进行训练,分别获得对应图像失真类型和对应评价算法的多通道融合评价结果分值。

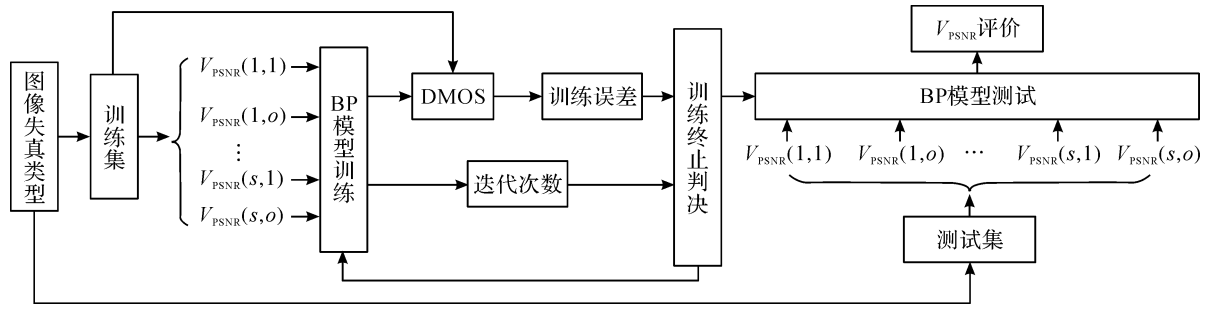


图 2 多通道评价融合模型

Fig. 2 Pooling model of multi-channel assessment

1.3 自适应融合原理

根据 V_{PSNR} 、 V_{SSIM} 、 V_{GSM} 和 V_{SVD} 的评价结果,基于回归函数将其进行自适应融合^[14],得最终结果 V_{VPL} 分值如下:

$$V_{\text{VPL}} = [(V_{\text{PSNR}}^{1-\lambda_1} V_{\text{SSIM}}^{\lambda_1})^{1-\lambda_2} V_{\text{GSM}}^{\lambda_2}]^{1-\lambda_3} V_{\text{SVD}}^{\lambda_3} \quad (7)$$

式(7)中: $\lambda_1 = 1/[1 + \gamma_{11} V_{\text{SSIM}}^{\gamma_{12}}]$, $\lambda_2 = 1/[1 + \gamma_{21} V_{\text{GSM}}^{\gamma_{22}}]$, $\lambda_3 = 1/[1 + \gamma_{31} V_{\text{SVD}}^{\gamma_{32}}]$, 参数 γ_{11} 、 γ_{12} 、 γ_{21} 、 γ_{22} 、 γ_{31} 和 γ_{32} 根据经验和试验训练取值,参见表 1。

2 结果分析

2.1 VPL 方法结果分析

基于 LIVE(database release2)数据库的试验结果如表 1 所示,试验挑选了近年来的主流方法进行比较,其中,特征相似度(feature similarity, FSIM)^[6]和视觉显著索引(visual saliency index, VSI)^[7]基于视觉特征评价原理,信息保真度(information fidelity criterion, IFC)^[8]和视觉信息保真(visual information fidelity, VIF)^[9]基于信息评价原理,显著失真(most apparent distortion, MAD)^[15]和内推机制(internal generative mechanism, IGM)^[16]基于视觉心理评价原理。表 1 的试验结果表明本研究 VPL 方法的优势,体现在:1)所有失真类型评价的均方根误差(root mean square error, RMSE)和皮尔逊线性相关系数(Pearson linear correlation coefficient, PLCC)两项指标均具有较高的水平, RMSE 小于 3.663 4, PLCC 大于 0.988 3, 明显优于其他方法。2)斯皮尔曼阶相关系数(Spearman rank-order correlation coefficient, SROCC)指标在 JPEG2000、WN、gblur 和 fastfading 多个失真类型上均有较大的优势。3)本方法的 3 项指标对各种失真类型评价都保持了非常稳定的高水平, RMSE 小于等于 3.663 4, PLCC 大于等于 0.988 3, SROCC 大于等于 0.960 0, 没有出现明显的下降,而其他方法均有一定程度的下降。4)对各种失真类型的总体(All)评价上,本方法的 3 项指标水平具有明显优势, RMSE 为 4.421 0, PLCC 为 0.983 7, SROCC 为 0.981 5。综上,本研究 VPL 方法相比其他方法具有较高的准确度和稳定性。

表 1 VPL 法与其他方法的评价结果比较

Table 1 Comparison of assessment results between VPL and other methods

失真类型	指标	IFC	VIF	GSM	FSIM	VSI	SVD	IGM	MAD	VPL
JPEG2000 $\gamma_{11}=2.50$ $\gamma_{12}=0.2$ $\gamma_{21}=30.0$ $\gamma_{22}=0.000\ 1$ $\gamma_{31}=1.0$ $\gamma_{32}=0.6$	RMSE	4.696 1	4.714 8	4.179 3	4.408 6	5.496 8	4.445 8	3.310 5	3.059 1	2.846 7
	PLCC	0.981 1	0.980 9	0.985 1	0.983 4	0.974 0	0.983 1	0.990 7	0.992 0	0.993 1
	SROCC	0.964 1	0.978 7	0.986 9	0.988 4	0.969 8	0.981 4	0.986 6	0.987 1	0.988 0
JPEG $\gamma_{11}=0.34$ $\gamma_{12}=10.0$ $\gamma_{21}=0.30$ $\gamma_{22}=6.0$ $\gamma_{31}=0.01$ $\gamma_{32}=1.6$	RMSE	3.251 1	2.802 7	3.909 6	3.861 5	3.941 4	3.419 4	2.336 6	4.532 8	2.377 6
	PLCC	0.991 3	0.993 6	0.987 4	0.987 7	0.987 2	0.990 4	0.995 5	0.983 0	0.995 4
	SROCC	0.953 8	0.965 0	0.965 0	0.980 1	0.954 1	0.953 6	0.976 0	0.967 1	0.960 0
WN $\gamma_{11}=0.1$ $\gamma_{12}=0.3$ $\gamma_{21}=0.000\ 1$ $\gamma_{22}=0.1$ $\gamma_{31}=40.0$ $\gamma_{32}=0.000\ 1$	RMSE	3.078 6	4.131 3	3.307 9	3.933 6	3.335 2	2.374 2	1.659 0	3.124 3	1.278 3
	PLCC	0.991 7	0.985 1	0.990 5	0.986 5	0.990 3	0.995 1	0.997 6	0.991 5	0.998 6
	SROCC	0.957 7	0.966 6	0.987 9	0.976 9	0.974 7	0.994 0	0.994 5	0.991 3	0.992 2
Gblur $\gamma_{11}=0.14$ $\gamma_{12}=0.1$ $\gamma_{21}=0.003$ $\gamma_{22}=1.6$ $\gamma_{31}=20.0$ $\gamma_{32}=8.9$	RMSE	3.859 5	7.503 8	5.505 4	4.566 3	5.523 8	6.776 2	3.629 6	3.893 0	2.636 4
	PLCC	0.988 2	0.954 6	0.975 8	0.983 4	0.975 7	0.963 1	0.989 6	0.988 0	0.994 5
	SROCC	0.969 0	0.933 4	0.987 9	0.991 6	0.967 9	0.963 5	0.988 8	0.987 7	0.992 0
Fastfading $\gamma_{11}=0.1$ $\gamma_{12}=0.3$ $\gamma_{21}=0.001$ $\gamma_{22}=1.8$ $\gamma_{31}=10.0$ $\gamma_{32}=0.6$	RMSE	5.117 4	4.125 8	5.078 8	5.258 1	5.553 9	6.184 3	5.086 5	4.192 0	3.644 9
	PLCC	0.977 1	0.985 2	0.977 5	0.975 8	0.973 0	0.966 4	0.977 4	0.984 7	0.988 5
	SROCC	0.957 0	0.969 5	0.982 7	0.981 5	0.964 4	0.969 9	0.972 0	0.985 0	0.985 0
All $\gamma_{11}=0.1$ $\gamma_{12}=0.3$ $\gamma_{21}=0.001$ $\gamma_{22}=1.8$ $\gamma_{31}=10.0$ $\gamma_{32}=0.6$	RMSE	5.371 8	7.469 2	6.103 1	5.784 3	6.597 9	6.279 0	4.722 0	5.517 2	4.421 0
	PLCC	0.975 8	0.952 6	0.968 6	0.971 9	0.963 2	0.966 8	0.981 3	0.974 4	0.983 7
	SROCC	0.955 2	0.920 6	0.971 8	0.977 8	0.949 1	0.967 2	0.974 1	0.977 3	0.981 5

2.2 算法复杂度测试

本试验测试了上述方法的运行时间,结果如表 2 所示,测试平台为 i7-7500U CPU@2.70 GHz、8 G 内存的 L470 便携式 PC、Windows10、MATLAB R2014a。表 2 中数据说明,与基于视觉神经理论的同类方法 IGM、MAD 相比,本研究 VPL 方法的运行时间与 MAD 基本上相当,但明显优于 IGM 方法。

图 3(a)~(d)分别为 IFC、FSIM、VSI 和 VPL 方法的散点分布图,其中,VPL 方法的客观评价结果均匀地分布在 $[0,100]$ 区内,这表明它具有最高的分辨率和线性度。

表 2 不同算法运行时间比较

Table 2 Comparison of running time among different algorithms

评价方法	时间/s	评价方法	时间/s
IFC	1 228	SVD	300
VIF	1 259	IGM	19 275
GSM	78	MAD	2 244
FSIM	450	VPL	2 710
VSI	253		

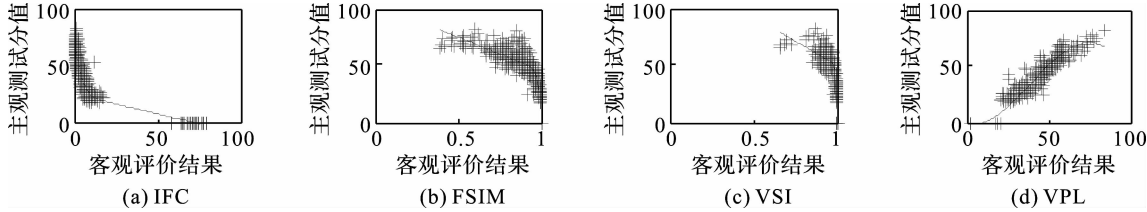


图 3 不同方法散点分布图

Fig.3 Scatter distribution plot of different methods

3 结 语

针对 PSRN、SSIM、GSM 和 SVD 主流评价方法的不足,本研究通过引入 HVS 的显著感知特性对其进行了改进,并借助于 BP 神经网络的学习推理能力对 HVS 的多通道评价进行了融合,解决了 HVS 中

多通道评价融合的不稳定性问题。试验结果表明我们所提方法具有较高的准确度和稳定性,因而具有一定的参考价值。

参考文献:

- [1] GUO Y C, HAO Y T, YU M. Image retargeting quality assessment based on content deformation measurement[J]. Signal Processing:Image Communication,2018,67(6):171.
- [2] ZHANG Y B, LIN W S, LI Q H, et al. Multiple-level feature-based measure for retargeted image quality[J]. IEEE Transactions on Image Processing,2018,27(1):451.
- [3] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing,2004,13(4):600.
- [4] LIU A M, LIN W S, NARWARIA M. Image quality assessment based on gradient similarity[J]. IEEE Transaction on Image Processing,2012,21(4):1500.
- [5] SHNAYDERMAN A, GUSEV A, ESKICIOGLU A M. An SVD-based grayscale image quality measure for local and global assessment[J]. IEEE Transactions on Image Processing,2006,15(2):422.
- [6] ZHANG L, ZHANG L, MOU X Q, et al. FSIM: a feature similarity index for image quality assessment[J]. IEEE Transaction on Image Processing,2011,20(8):2378-2386.
- [7] ZHANG L, SHEN Y, LI H Y. VSI: a visual saliency induced index for perceptual image quality assessment[J]. IEEE Transactions on Image Processing,2014,23(10):4270.
- [8] SHEIKH H R, BOVIK A C, DE VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. IEEE Transactions on Image Processing,2005,14(12):2117.
- [9] SHEIKH H R, BOVIK A C. Image information and visual quality[J]. IEEE Transactions on Image Processing, 2006,15(2):430.
- [10] 丰明坤,赵生妹,邢超. 一种基于视觉特性的 PSNR 图像评价方法[J]. 南京邮电大学学报(自然科学版),2015,35(8):35.
- [11] 丰明坤,赵生妹,邢超. 基于视觉显著失真度的图像质量自适应评价方法[J]. 电子与信息学报,2015,37(9):2064.
- [12] 丰明坤,赵生妹,施祥. 视觉多通道梯度与低阶矩自适应图像评价[J]. 仪器仪表学报,2015,36(11):2533.
- [13] 丰明坤,王中鹏,叶绿. 视觉稀疏化多通道多特征自适应的图像评价[J]. 仪器仪表学报,2016,37(3):669.
- [14] 丰明坤,施祥. 视觉特征深度融合的图像质量评价[J]. 浙江大学学报(工学版),2019,53(3):516.
- [15] LARSON E C, CHANDLER D M. Most apparent distortion: full-reference image quality assessment and the role of strategy[J]. Journal of Electronic Imaging,2010,19(1):011006-1.
- [16] WU J J, LIN W S, SHI G M, et al. Perceptual quality metric with internal generative mechanism[J]. IEEE Transactions on Image Processing,2013,22(1):43.